



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07H 21/04, C07K 1/04, 1/14, 1/36, 2/00, 4/00, 5/02, 7/02	A1	(11) International Publication Number: WO 98/08865 (43) International Publication Date: 5 March 1998 (05.03.98)
(21) International Application Number: PCT/US97/15295 (22) International Filing Date: 2 September 1997 (02.09.97) (30) Priority Data: 60/024,881 30 August 1996 (30.08.96) US 08/920,874 29 August 1997 (29.08.97) US (71) Applicant: LIFE TECHNOLOGIES, INC. [US/US]; 9800 Medical Center Drive, Rockville, MD 20850 (US). (71)(72) Applicants and Inventors: GOSS, Dixie, J. [US/US]; Apartment 19M, 180 West End Avenue, New York, NY 10023 (US). FRIEDLAND, Diana [US/US]; 695 Park Avenue, New York, NY 10023-5024 (US). (74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: CAP-BINDING DOMAIN OF HUMAN EUKARYOTIC PROTEIN SYNTHESIS INITIATION FACTOR eIF-4E AND THE USE THEREOF		
(57) Abstract Binding of eIF-4E to the 5'm⁷G cap structure of eukaryotic mRNA signals the initiation of protein synthesis. In the present invention, the binding site cap-binding domain of eIF-4E was identified as the region containing the sequence Trp¹¹³ to Arg¹²². Thus, in a first preferred embodiment the present invention provides a truncated peptide of the cap binding protein consisting essentially of the mRNA cap-binding domain. Specifically, the invention concerns eIF-4E cap binding peptide and its variants. In additional preferred embodiments, the invention provides an isolated nucleic acid molecule which encodes a truncated peptide of the eIF-4E protein consisting essentially of the mRNA cap-binding domain of eIF-4E; a recombinant vector and host cell comprising the above-described nucleic acid molecule; and a recombinant peptide produced thereby. The invention further provides the truncated or recombinant peptides immobilized on a solid support, a resin comprising the truncated or recombinant peptides immobilized on a solid support, methods and kits for isolating a capped mRNA molecule, and methods and kits for producing a full-length cDNA molecule.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**Cap-Binding Domain of Human
Eukaryotic Protein Synthesis Initiation Factor eIF-4E
and the Use Thereof**

STATEMENT AS TO U.S. GOVERNMENT RIGHTS

5 The present invention was made in part under grants from the National Science Foundation (GER-9023681, MCB 9303661 and DGE-9553549) and the National Institutes of Health (NIH-CA63640). The U.S. government has certain rights in the present invention.

FIELD OF THE INVENTION

10 The present invention is in the field of molecular biology. Specifically, the invention provides truncated peptides of the eukaryotic protein synthesis initiation factor eIF-4E comprising the region of eIF-4E that binds to the cap regions of mRNA molecules. The invention also provides methods and kits using these peptides suitable for use in the isolation of capped mRNA molecules and in the production of full-length cDNA molecules.

RELATED ART

15 **Overview**

20 The initiation phase of eukaryotic protein synthesis is characterized by recognition of the m⁷(5')Gppp(5')N (where "N" is any nucleotide) cap structure at the 5' terminus of eukaryotic mRNA by initiation factors (eIF's) (reviewed in Rhoads, R.E., *Trends Biochem. Sci. (Pers. Ed.)* 13:52-56 (1988); Sonenberg, N., *Prog. Nucleic Acid Res. Mol. Biol.* 35:173-207 (1988)). Of the eIF's which interact at or near the cap, only eIF-4E (alone or as part of the eIF-4F complex) has been shown to interact directly with the cap (Sonenberg, N., *Nucleic Acids Res.* 9:1643-1656 (1981); Sonenberg, N., *et al.*, *Cell* 27:563-572 (1981); Tahara, S.M., *et al.*, *J. Biol. Chem.* 256:7691-694 (1981); Hellmann, G.M., *et al.*, *J. Biol. Chem.* 257:4056-4062 (1982); Grifo, H.A., *et al.*, *J. Biol. Chem.* 258:5804-5810 (1981); Webb, N.R., *et al.*, *Biochemistry* 23:177-181

(1984)). This initial recognition of the cap structure represents the first committed step in the initiation phase of protein synthesis.

Evidence that cap recognition by eIF-4E is a limiting step in protein synthesis is supported by the observation that it is the least abundant of all the initiation factors and thus may serve a regulatory function (Duncan, R.F., *et al.*, *J. Biol. Chem.* 262:380-388 (1987); Hiremath, L.S., *et al.*, *J. Biol. Chem.* 260:7843-7849 (1985)). In addition, there are a number of other observations indicating that this protein is a key point for post-transcriptional control of gene expression. For example, EGF and other growth regulatory peptides rapidly stimulate the phosphorylation of eIF-4E phosphorylation in intact cells (Donaldson, R.W., *et al.*, *J. Biol. Chem.* 266:3162-3166 (1991); Sonenberg, N., *Biochemie* 76:839-846 (1994); Rhoads, R.E., *J. Biol. Chem.* 268:3017-3020 (1993); Proud, C.G., *Curr. Top. Cell. Regul.* 32:243-369 (1992)). Furthermore, evidence has been obtained from two different cellular systems that hyperphosphorylation of eIF-4E promotes the recruitment of other eIF-4E subunits to the m⁷G mRNA cap (Bu, X., *et al.*, *J. Biol. Chem.* 268:4975-4978 (1993); Morley, S.J., *et al.*, *Eur. J. Biochem.* 218:39-48 (1993)) and a three- to four-fold increase in the affinity of eIF-4E binding to mRNA caps (Minich, W.B., *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 91:7668-7672 (1994)). The regulation of eIF-4E function has been shown to be even more complicated by the recent identification of eIF-4E binding proteins that are regulated physiologically by insulin (Lin, T.A., *et al.*, *Science* 266:653-656 (1994)). When bound to eIF-4E these proteins prevent its recognition of, and binding to, m⁷G caps of mRNA. To understand fully the mechanisms by which these covalent modifications and translational repressor 4E binding proteins exert their effects a more detailed understanding of eIF-4E structure and binding site sequence will be required.

Structure of the eIF-4E Cap-binding Domain

The amino acid sequence for the cap-binding protein has been previously reported for several species including yeast (Altmann, M., *et al.*, *Mol. Cell. Biol.* 7:998-1003 (1987)), wheat (Metz, A.M., *et al.*, *Nucl. Acids Res.* 20:4096 (1992)), human (Rychlik, W., *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:945-949 (1987)), and mouse (Altmann, M., *et al.*, *J. Biol. Chem.*

264:12145-12147 (1989)). While there is some interspecies divergence in the amino acid sequences of the cap-binding proteins, the numbers and positions of eight tryptophan residues are highly conserved, with there being one additional tryptophan residue in p26 of wheat germ eIF-4F (Metz, A.M., *et al.*, *Nucl. Acids Res.* 20:4096 (1992)).

5 Although no x-ray crystallographic or structural NMR information on the eIF-4E•cap complex have yet been reported, photoaffinity labeling studies of mammalian eIF-4E have suggested that the cap-binding domain is within amino acid residues 47-182 of the native protein (Chavan, A.J., *et al.*, *Biochemistry* 29: 5521-5529). Furthermore, mutational analysis suggested that a glutamic acid residue, which is located three amino acid residues to the carboxy side of
10 tryptophan 5 (numbering of tryptophan residues proceeds from the amino terminus to the carboxy terminus), participates in the binding of the mRNA cap to human eIF-4E via direct hydrogen bond formation with the 2-amino group of the m⁷G base (Ueda, H., *et al.*, *FEBS Lett.* 280: 207-210 (1991); Ueda, H., *et al.*, *Biochem. Biophys. Acta.* 1075:181-186 (1991)). Interestingly, however, previous work from this same group had proposed that a glutamic acid residue two amino acid
15 residues to the carboxy side of tryptophan 7 participates in this hydrogen bond interaction (Ueda, H., *et al.*, *Biochem. Biophys. Res. Commun.* 154:199-204 (1988)). Regardless of which conclusion is correct, however, the cap-binding subunit of wheat germ eIF-4F lacks any such acidic residue in these regions, containing instead a β -turn-inducing proline residue in the region of tryptophan 5 and a basic histidine residue in the region of tryptophan 7 (Metz, A.M., *et al.*,
20 *Nucleic Acids Res.* 20:4096 (1992)). These conflicting results and the lack of structural information indicate the need for a more precise localization of the eIF-4E binding site for the 5' m⁷G cap of mRNAs.

Photoaffinity Labeling

25 Photoaffinity labeling of nucleotide binding sites has provided a highly selective method for identifying the participating amino acids of many proteins (Salvucci, M.E., *et al.*, *Biochemistry* 31:4479-4487 (1992); Shoemaker, M.T., and Haley, B.E., *Biochemistry* 32:1883-1890 (1993); Jayaram, B., and Haley, B.E., *J. Biol. Chem.* 269:3233-3242 (1994)). Photoactivation of the azido moiety with ultra-violet light results in the generation of a nitrene and the subsequent

covalent attachment of the probe to the binding domain through the remaining nitrogen (Potter, R.L., and Haley, B.E., *Meth. Enzymol.* 91:613-633 (1983)). Immobilized aluminum(III)-chelate chromatography has facilitated the isolation of binding site peptides due to the presence of highly charged phosphates on the photoinserted azidonucleotide which interact with the Al^{3+} (Salvucci, M.E., *et al.*, *Biochemistry* 31:4479-4487 (1992); Shoemaker, M.T., and Haley, B.E., *Biochemistry* 32:1883-1890 (1993); Jayaram, B., and Haley, B.E., *J. Biol. Chem.* 269:3233-3242 (1994); Anderssen, L. J., *Chromatography* 539: 327-334 (1991)).

Uses of eIF-4E Protein

Isolated cap-binding eIF-4E proteins have several uses. For example, U.S. Patent No. 5,219,989 to Sonenberg *et al.* discloses the preparation of multifunctional fusion proteins comprising the eIF-4E protein and at least a second protein having the ability to bind to a solid support matrix. The exemplified second protein is protein A which is capable of binding to a resin having an IgG antibody attached to it. Such fusion proteins may be used for the preparation of a cDNA library mostly containing full-length cDNA and for the purification of capped mRNA.

A preparation consisting essentially of the cap-binding region of eIF-4E (or fragments thereof), however, has not been reported. Such a preparation would provide a less expensive and more specific means for binding of capped mRNA molecules, and for production of full-length cDNA molecules, than the fusion proteins of U.S. Patent No. 5,219,989. In addition, such a preparation could be produced more efficiently in recombinant systems such as mammalian host cells, since the nucleic acid sequences directing the synthesis of such a preparation would be smaller and therefore more efficiently incorporated and transcribed than those for fusion proteins.

It is therefore an object of the present invention to provide the amino acid sequence of the cap-binding region of eIF-4E, to provide peptides and fragments thereof comprising this region, and to provide methods for the use of these peptides and fragments in the isolation of capped mRNA molecules and in the preparation of full-length cDNA molecules.

BRIEF SUMMARY OF THE INVENTION

The present invention provides a truncated peptide of the eIF-4E protein consisting essentially of the mRNA cap-binding domain of eIF-4E, wherein the peptide has an amino acid sequence corresponding to that set forth in SEQ ID NO:1. The invention further encompasses
5 variant peptides comprising the cap-binding domain of eIF-4E within a β -sheet structural motif. In preferred embodiments, the invention provides the peptides immobilized on a solid support, a resin comprising the peptides immobilized on a solid support, a method for isolating a capped mRNA molecule, a method for producing a full-length cDNA molecule, a kit for isolating a capped mRNA molecule and a kit for producing a full-length cDNA molecule.

10 Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of the following drawings and description of the invention, and of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A is a graph and an autoradiogram of a gel demonstrating saturation of $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ photoincorporation into eIF-4E. eIF-4E (4 micrograms) was incubated with the indicated
15 concentrations of $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ in 50 μl of photolysis buffer. The reaction mixture was irradiated with UV light for 90 seconds and analyzed by SDS-PAGE. ^{32}P incorporation was determined by radioisotopic imaging and quantification.

FIG. 1B is a graph and an autoradiogram of a gel demonstrating competitive inhibition of $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ photoincorporation into eIF-4E by m^7GTP . eIF-4E (4 micrograms) was
20 incubated with 40 μM $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ in the presence of increasing concentrations of m^7GTP under the conditions described for Fig. 1A. ^{32}P incorporation was determined by radioisotopic imaging and quantitation.

FIG. 2 is a bar graph representing a chromatogram (radioactivity profile) of immobilized Al^{3+} chromatography of tryptic peptides. eIF-4E (5 mg) was photolabeled with 50 μM $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$
25 followed by a second photolabeling with 50 μM $\text{8-N}_3\text{GTP}$. Following digestion with modified trypsin, Al^{3+} chelate chromatography proceeded as described in the Example. Fractions

1-5 represent flow through, 6-25 represent various buffer washes, and 26-31 represent fractions eluted with 10 mM K_2HPO_4 . ^{32}P levels were determined by liquid scintillation counting.

FIG. 3A is a chromatogram of microbore, C_8 reverse phase HPLC of tryptic peptide fractions from Al^{3+} chelate chromatography, demonstrating a typical UV profile of the radioactive fractions resulting from K_2HPO_4 elution of the Al^{3+} chelate column.

FIG. 3B is a chromatogram demonstrating the corresponding ^{32}P cpm profile of the fractions resulting from the HPLC in Fig. 3A. Radioactivity levels were determined by liquid scintillation counting.

DETAILED DESCRIPTION OF THE INVENTION

10 Identification of the Cap-binding Domain of eIF-4E

Binding studies of m^7GTP , m^7GpppG , and various capped mRNAs with eIF-4E have provided insight into the nature of the eIF-4E•cap interaction (Carberry, S.E., *et al.*, *Biochemistry* 28:8078-8083 (1989); Carberry, S.E., *et al.*, *Biochemistry* 31:1427-1432 (1992); Goss, D.J., *et al.*, *Biochem. Biophys. Acta.* 1050:163-166 (1990)). Tryptophan stacking and hydrogen bonding with glutamic acid residues have been suggested to be important factors in binding of the mRNA cap structure by native eIF-4E (Kamiichi, K., *et al.*, *J. Chem. Soc. Perkin Trans. II*, 1739-1745 (1987); Ishida, T., *et al.*, *J. Am. Chem. Soc.* 110:2286-2294 (1988); Ueda, H., *et al.*, *Biochem. Biophys. Acta.* 1075:181-186 (1991)). Additionally, the involvement of a protonated histidine in cap binding by eIF-4E has been proposed (Carberry, S.E., *et al.*, *Biochemistry* 28:8078-8083 (1989)). These results have allowed for much speculation regarding the specific amino acid residues that constitute the m^7G cap binding site.

In the examples of the present invention provided below, the photoaffinity probe [γ - ^{32}P]8- N_3GTP was used as a photoactivable analogue of the m^7G cap structure of mRNA to selectively label the m^7GTP cap-binding domain of human recombinant eIF-4E (eIF-4E). The photolabeled peptide resulting from proteolysis was isolated using aluminum(III)-chelate chromatography followed by reversed-phase high performance liquid chromatography. The amino acid sequence of this peptide has been identified in the present invention as consisting essentially of residues 113-

122 of the amino acid sequence of human eIF-4E described previously (Rychlik, W., *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:945-949 (1987)). This peptide thus contains tryptophan 6 (Trp¹¹³ in full-length eIF-4E).

Amino Acid sequence of the Cap-binding Domain of eIF-4E

5 The present invention is thus related to a truncated cap binding protein (e.g. eIF-4E protein) which is capable of binding to the 5' terminus cap structure of mRNA produced in eukaryotic cells. Such truncated proteins comprise at least amino acids Trp¹¹³ to Arg¹²² corresponding to the human eIF-4E sequence as disclosed by Rychlik, W., *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:945-949 (1987), or suitable variants thereof. Examples of such truncated
10 proteins include:

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg (SEQ ID NO: 1).

Variants of the truncated eIF-4E protein include any change or changes (deletions, substitutions or additions) in or to the amino acid sequence, provided that the variant retains the ability to bind to the cap region of mRNA. The ability to retain cap binding function can be
15 determined by the assays described herein. Preferably, changes are made so as to conserve the hydrophobicity and charge of the peptide. Examples of such variants include but are not limited to one or more substitutions within the peptide as follows:

20 Leu changed to Ala, Val, Ile or Phe;
 Asn changed to Gln;
 Gln changed to Asn;
 Thr changed to Ser or Gly;
 Lys changed to Arg or His; and
 Arg changed to Lys or His.

Previous fluorescence quenching studies have indicated the involvement of a tryptophan residue, which apparently resides in a somewhat hydrophobic environment, in eIF-4E binding of the mRNA cap (Carberry, S.E., *et al.*, *Biochemistry* 28:8078-8083 (1989)). Computer-predicted secondary structure obtained using Chou and Fasman sequence analysis (Chou, P.Y., and Fasman, G.D., *Biochemistry* 13: 222-245 (1974)) indicate that tryptophan 6 comes after a putative turn region in the protein and would be at least partially buried within a hydrophobic core with only partial exposure to the aqueous environment (McCubbin, W.D., *et al.*, *J. Biol. Chem.* 263:17663-17671 (1988)). The calculations also suggest that tryptophan 6 lies within a region of β -sheet structure. Circular dichroism data obtained in studies related to those disclosed here reveal that a large conformational transition occurs in wheat eIF-(iso)4F upon its binding of mRNA cap analogues. Taken together, these data suggest that the interaction between eIF-4E and the mRNA cap may involve a structural β -sheet motif; this conformational transition may thus play a regulatory role, as has been proposed for other proteins containing similar motifs (Wang, Y., *et al.*, in: *Spectroscopy of Biological Molecules*, Merlin, J.C., *et al.*, eds.), pp. 335-336 (1995)), by determining the interaction of the eIF-4E binding domain with the m⁷G cap structure of mRNA.

It is well-known in the art that β -strands, which are organized into higher secondary structures such as β -sheets and α/β barrels in a variety of proteins, share certain general properties with respect to their preferred amino acid sequences (Sternberg, M.J.E., *et al.*, *Phil. Trans. R. Soc. Lond. B* 293:177-189 (1981); Wodak, S.J., *et al.*, *Biochem. Soc. Symp.* 57:99-121 (1990); Siezen, R.J., *et al.*, *Prot. Eng.* 4(7):719-737 (1991); Chothia, C., *et al.*, *Ciba Found. Symp.* 162:36-57 (1991); Bork, P., *et al.*, *J. Mol. Biol.* 242:309-320 (1994); Yoshida, M., and Amano, T., *FEBS Letts.* 359:1-5 (1995)). In general, β -strands comprising β -sheet structural motifs in most proteins are composed of stretches of nonpolar or hydrophobic amino acid residues and are typically 3-10 residues in length. While any hydrophobic residue may be found in these regions, including alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), phenylalanine (Phe), methionine (Met), tryptophan (Trp), proline (Pro) or cysteine (Cys), the branched aliphatic residues (Leu, Ile and Val) are particularly preferred (Sternberg, M.J.E., *et al.*, *Phil. Trans. R. Soc. Lond. B* 293:177-189 (1981)). In addition, β -sheet motifs typically have one or more charged (especially aspartic acid (Asp) or glutamic acid (Glu)), aromatic (especially tyrosine (Tyr))

or threonine (Thr)) or β -turn-inducing (especially Pro or glycine (Gly)) residues immediately prior to and/or following the stretch of hydrophobic amino acids (Sternberg, M.J.E., *et al.*, *Phil. Trans. R. Soc. Lond. B* 293:177-189 (1981); Wodak, S.J., *et al.*, *Biochem. Soc. Symp.* 57:99-121 (1990); Siezen, R.J., *et al.*, *Prot. Eng.* 4(7):719-737 (1991); Chothia, C., *et al.*, *Ciba Found. Symp.* 162:36-57 (1991); Bork, P., *et al.*, *J. Mol. Biol.* 242:309-320 (1994); Yoshida, M., and Amano, T., *FEBS Letts.* 359:1-5 (1995)). The amino acid residues thus arranged in the primary structure of the protein or peptide interact via aromatic ring stacking, hydrogen bonding and hydrophobic interactions to form β -strands and the resulting higher order secondary structures such as β -sheets.

10 Accordingly, the present invention also encompasses variants of the eIF-4E peptide described above (for SEQ ID NO:1) in which Trp113 (tryptophan 6) is maintained in the same β -sheet structural motif. Included in these variants are peptides having an amino acid sequence corresponding to the general pattern

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X₁₍₁₋₂₎-X₂₍₄₋₁₀₎-X₁₍₁₋₂₎ (SEQ ID NO:2)

15 wherein X₁ is a charged, aromatic or β -turn-inducing amino acid, most preferably Asp, Glu, Tyr, Thr, Gly or Pro; X₂ is a hydrophobic amino acid, most preferably Leu, Ile, Val or Trp; and the subscripts for X₁ and X₂ indicate the number of such residues in consecutive sequence at the indicated position.

Other preferred variant peptides of the invention include those having an amino acid
20 sequence corresponding to the general pattern

X_n-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X_n (SEQ ID NO:3)

wherein X is any amino acid, and n is a whole number ranging in value from 0 to 70, preferably 0 to 50, 0 to 25 or 0 to 10.

Other preferred variant peptides of the invention include those having an amino acid
25 sequence corresponding to the general pattern

X_n -Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg- X_n (SEQ ID NO:4)

wherein X is any amino acid, and n is a whole number ranging in value from 0 to 70, preferably 0 to 50, 0 to 25 or 0 to 10.

Other preferred variant peptides of the invention include those having an amino acid
5 sequence corresponding to the general pattern

X_n -Trp-Glu-Asp-Glu-Lys-Asn-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg- X_n (SEQ ID NO:5)

wherein X is any amino acid, and n is a whole number ranging in value from 0 to 70, preferably 0 to 50, 0 to 25 or 0 to 10.

10 Other preferred variant peptides of the invention include those having an amino acid sequence corresponding to the general pattern

$X1_{(0-10)}-X2_{(0-2)}-X1_{(0-10)}$ -Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-
 $X1_{(1-2)}-X2_{(4-10)}-X1_{(1-2)}$ (SEQ ID NO:6)

15 wherein X1 is generally a charged, aromatic or β -turn-inducing amino acid (such as Asp, Glu, Tyr, Thr, Gly or Pro), X2 is a hydrophobic amino acid (such as Leu, Ile, Val or Trp) and the subscripts for X1 and X2 indicate the number of such residues in consecutive sequence at the indicated position.

Examples of such variant peptides include, but are not limited to:

Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg (SEQ ID NO:7);

20 Trp-Glu-Asp-Glu-Lys-Asn-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Lys-Lys (SEQ ID NO:8);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Gly-Asp-Ile-Val-Val-Leu-Val-Pro (SEQ ID NO:9);

-11-

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Asp-Gly-Val-Val-Ile-Ile-Val-Leu-Leu-Asp-Gly (SEQ ID NO:10);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Asp-Leu-Leu-Val-Ile-Val-Glu-Gly (SEQ ID NO:11);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Glu-Asp-Val-Ile-Val-Leu-Leu-Ile-Tyr (SEQ ID NO:12);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Pro-Ile-Val-Val-Ile-Val-Leu-Val-Ile-Ile-Asp-Gly (SEQ ID NO:13);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Asp-Thr-Val-Leu-Val-Val-Ile-Gly-Pro (SEQ ID NO:14);

10 Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Gly-Asp-Ile-Ile-Val-Val-Val-Ile-Val-Asp-Pro (SEQ ID NO:15);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Gly-Val-Val-Val-Ile-Ile-Val-Leu-Leu-Asp (SEQ ID NO:16);

Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Pro-Thr-Val-Ile-Val-Val-Ile-Asp (SEQ ID NO:17);

15 Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg (SEQ ID NO:18);

Trp-Glu-Asp-Glu-Lys-Asn-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Lys-Lys (SEQ ID NO:19);

Lys-Leu-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Val-Leu-Gln (SEQ ID NO:20);

Ile-Gln-Leu-Val-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Leu-Lys-Ala-Gly (SEQ ID NO:21);

Val-Ala-Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Gln-Ala-Lys-Thr (SEQ ID NO:22);

Ile-Thr-Val-Gln-Ala-Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Lys-
5 Thr-Leu-Phe (SEQ ID NO:23);

Leu-Arg-Val-Trp-Glu-Asp-Glu-Lys-Asn-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Lys-Tyr (SEQ ID NO:24);

Asn-Ile-Ser-Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Phe-Lys-Ala-Val-Gln-Gln (SEQ ID NO:25);

10 Glu-Tyr-Glu-Gly-Val-Val-Pro-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gly-Glu-Arg-Asp-Leu-Ile-Leu-Val-Ile-Gly (SEQ ID NO:26); and

Asp-Gly-Pro-Gly-Leu-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-Tyr-Thr-Trp-Val-Val-Ile-Leu-Val-Thr-Gly (SEQ ID NO:27).

Other amino acid sequences for peptides that may be used in the present invention will be
15 apparent to one of ordinary skill in the art, in light of the above descriptions of preferred sequences. Based on structural and thermodynamic considerations as described above, this general sequence of amino acids should localize Trp113 within the β -strand or β -sheet structural motif that is optimal for maintenance of mRNA cap recognition and binding by eIF-4E.

Synthesis of eIF-4E Peptides

20 Truncated eIF-4E peptides such as those described above may be prepared according to well-known methods of solid phase peptide synthesis, using routine methods of organic and

inorganic chemistry. Alternatively, the peptides may be produced by recombinant DNA techniques, which would yield a recombinant peptide comprising the above-described truncated eIF-4E peptides. Production of such recombinant peptides is preferably accomplished by expression of the peptide in a host cell, including a bacterial, yeast or mammalian cell, in which
5 the peptide coding sequence is operably linked to a promoter sequence. Typically, the promotor-peptide encoding sequence is contained by a vector. Techniques for isolation of nucleic acids, insertion into a vector, transformation of a host cell and isolation of the recombinant protein from the host cell are well-known in the art (*see, for example, Sambrook, J., et al., eds., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1992).*).

10 Use of eIF-4E Peptides

Isolation of Capped mRNA Molecules

In one preferred embodiment of the invention, the peptides produced as described above may be covalently or non-covalently immobilized on a solid phase support. By "solid phase support" is intended any solid support to which a peptide can be immobilized. Such solid phase
15 supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Preferred are beads made of glass, latex or a magnetic material. Linkage of the peptide of the invention to a solid support can be accomplished by attaching one or both ends of the peptide to the support. Attachment may also be made at one or more internal sites in the
20 peptide. Multiple attachments (both internal and at the ends of the peptide) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulfhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments, addition of an affinity tag sequence to the peptide can be used such as GST (Smith, D.B. and
25 Johnson, K.S. *Gene* 67:31 (1988); polyhistidines (Hochuli, E. *et al., J. Chromatog.* 411:77 (1987)); or biotin. Such affinity tags may be used for the reversible attachment of the peptide to the support.

Such immobilized peptides are useful in chromatographic isolation of mRNA molecules by binding to the cap region of the mRNAs. In this scenario, a solution containing the capped mRNA molecules would be contacted, *en masse* or on a column, with the resin containing the immobilized cap-binding peptides of the present invention. After allowing binding of the mRNA molecules to the resin, unbound materials may be simply washed away. The mRNA molecules may then be rapidly removed from the resin or other solid phase support by contacting the resin or support with an elution buffer which may contain, for example, a competitive analog of the mRNA cap such as m⁷GDP (*see* U.S. Patent No. 5,219,989), a buffer that causes a change in the pH, and/or ionic strength of the solution.

Immobilized capped mRNA molecules may instead be used directly in an RT-PCR method according to methods that are well known.

Production of Full-length cDNA Molecules

The mRNA cap-binding peptides of the present invention can be used to obtain full-length cDNAs, as has been described for eIF-4E/protein A fusion proteins (U.S. Patent No. 5,219,989).

Production of cDNA molecules is effected by reverse transcription of a desired mRNA with a protein having reverse transcriptase activity, for example, according to U.S. Patent No. 5,244,797 with reverse transcriptase, or with a DNA polymerase having reverse transcriptase activity. Preferably, the reverse transcriptase is a truncated or mutated reverse transcriptase that lacks RNase H activities that gives a much higher proportion of full-length cDNA compared to the wild type reverse transcriptase. Such reverse transcriptases (SuperscriptTM and Superscript IITM, respectively) may be obtained from Life Technologies, Inc. (Gaithersburg, MD).

Preferred DNA polymerases having reverse transcriptase activity are thermostable enzymes. Examples of DNA polymerases having reverse transcriptase activity include the wild type and mutant *Tne* enzymes (*see* WO96/10640 and copending application entitled "Cloned DNA Polymerases from *Thermotoga* and Mutants thereof" by A. John Hughes and Deb K. Chatterjee, filed August 14, 1996); *Tth* enzyme (U.S. Patent Nos. 5,192,674, 5,413,926, 5,242,818 and WO91/09950); and *Taq* (U.S. Patent Nos. 5,474,920, 5,079,352, 4,889,818 and 5,352,600). *See also* U.S. Patent Nos. 5,407,800, 5,322,770 and 5,310,652.

An intermediate in the production of cDNA molecules via this process is a hybrid mRNA:cDNA molecule. Once a mixture containing mRNA:cDNA hybrids has been obtained after reverse transcription of mRNA, it is incubated with a single-stranded RNA specific nuclease, such as a T_1 or T_2 nuclease, and an endonuclease that specifically attacks the 3' adjacent phosphodiester-bound GpN. The naturally modified m^7G part of the cap structure will not be recognized by this enzyme. T_1 does not attack RNA that is hybridized to DNA.

If the reverse transcriptase copies the entire length of the mRNA, or if it falls short by a few nucleotides such that there is no unhybridized GpN residue in the corresponding mRNA, T_1 will not degrade the mRNA and, as a result, the cap structure will remain covalently bound to the mRNA:cDNA hybrid. If, however, cDNA synthesis was not complete, the single-stranded RNA specific nuclease will degrade unpaired RNA and remove the cap structure from the mRNA:cDNA hybrid.

Following nuclease treatment, the mRNA:cDNA hybrids are incubated with the truncated eIF-4E protein of the present invention, which may be immobilized on a solid phase support as described above. As a result of this incubation, only those mRNA:cDNA hybrids that have a covalently attached cap structure will bind to the protein of the present invention. By immobilizing the truncated eIF-4E to a solid phase support, all of the non-capped containing hybrids, or incomplete cDNAs, will wash away. The bound full-length capped mRNA:cDNA hybrids may then be released competitively by treatment with a cap analog such as m^7GDP as described in U.S. 5,219,989.

The resulting purified fraction contains only full length or near full length first strand cDNAs which then act as templates for second strand synthesis. The steps for completing the cDNA library are the same as those normally used by those of ordinary skill in the art.

By use of a reverse transcriptase lacking RNase H activity to reverse transcribe the mRNA together with the purification of the capped mRNA:cDNA hybrids after nuclease digestion, it is possible to obtain an unexpectedly high yield of full length cDNA, even when the DNA is long. Thus, the present invention is a major advance in the art.

The present invention is also directed to a kit for isolation of capped mRNA molecules. Kits according to this aspect of the invention comprise a carrier means having in close confinement therein one or more container means, each of which contain the reagents described

herein. For example, a first container means may contain a truncated eIF-4E peptide of the present invention, which may be bound to a solid phase as described above. A second container means may contain a competitive analogue of the cap region of mRNA molecules, which is preferably m⁷GDP.

5 Also encompassed within the present invention is a kit for the production of full-length cDNA molecules. Kits according to this aspect of the invention comprise a carrier means having in close confinement therein one or more container means, each of which contain the reagents described herein. For example, a first container means may contain the truncated eIF-4E protein of the present invention, which may be bound to a solid phase as described above. A second
10 container means may contain a nuclease, which is preferably T₁ nuclease. A third container means may contain a protein having reverse transcriptase activity, which is preferably a truncated or mutant reverse transcriptase lacking RNase H activity. A fourth container means may contain an elution buffer which may contain a competitive analogue of the cap region of mRNA molecules, which is preferably m⁷GDP. A fifth container means may contain one or more
15 nucleoside triphosphates and at least one buffer salt. Additional container means may contain an oligo(dT) primer or a control RNA. Two or more of the enzymes described above may alternatively be combined into a single container means.

The enzymes may be present in a buffered solution or in lyophilized form. Alternatively, the enzymes may be stabilized by drying in the presence of a sugar such as trehalose (U.S. patent
20 nos. 5,098,893 and 4,824,938) or acacia gum, pectin, carboxymethyl cellulose, carboxymethylhydroxyethyl cellulose, guar, carboxy guar, carboxymethylhydroxypropyl guar, laminaran, chitin, alginates or carrageenan.

It will be readily apparent to one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are obvious and
25 may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following example, which is included herewith for purposes of illustration only and is not intended to be limiting of the invention.

Examples

Example 1: Identification of Cap-binding Domain of eIF-4E

Materials and Methods

Materials. Sequencing grade modified trypsin was from Boehringer Mannheim Biochemica (Indianapolis, IN). Electrophoresis reagents were from Bio-Rad (Melville, NY). HPLC reagents were from E.M. Science (Gibbstown, NJ). All other reagents were obtained from Sigma Chemical Company (St Louis, MO) and were molecular biology grade unless otherwise noted.

Recombinant eIF-4E. Human recombinant eIF-4E was purified from *Escherichia coli* containing a T7 polymerase-driven vector which was described in detail elsewhere (Baker *et al.*, *J. Biol. Chem.* 267:11495-11499 (1992). See also Rychlik *et al.*, *Proc. Natl. Acad. Sci USA* 84:945-949 (1987)). Recombinant eIF-4E was purified from bacterial lysates by m⁷GTP Sepharose affinity chromatography followed by Mono Q FPLC (Pharmacia, Piscataway, NJ) (Haas D.W. & Hagedorn C.H., *Arch. Biochem. Biophys.* 284:84-89 (1991); Baker B.F., *et al.*, *J. Biol. Chem.* 267:11495-9 (1992); Bu, X. & Hagedorn, C.H. *FEBS Lett.* 301:15-18 (1992)). The FPLC purification step was performed in 50 mM HEPES (pH 8.0), 1 mM MgCl₂, 10% glycerol, 1 mM DTT and used a 0-500 mM gradient of NaCl. Fractions that contained greater than 95% pure recombinant eIF-4E were identified by measuring absorbance at 280 nm and SDS-PAGE analysis. The isoelectric point of the recombinant eIF-4E was identical to that of the dephosphorylated iso-species of eIF-4E isolated from cultured human cells (Bu, X. & Hagedorn, C.H. *FEBS Lett.* 301:15-18 (1992)). In addition to its ability to bind the m⁷GTP cap structure of mRNA, the recombinant eIF-4E prepared using these methods is phosphorylated *in vitro* by protein kinase C at the same sites as native eIF-4E (Haas D.W. & Hagedorn C.H., *Second Messengers and Phosphoproteins* 14:55-63 (1992)).

Synthesis of Photoaffinity Probe. The radioactive photo probe [γ -³²P]8-N₃GTP (specific activity, 32 mCi/ μ mol) was synthesized and purified as previously reported (Potter, R.L. & Haley, B.E. *Methods Enzymol.* 91:613-633 (1983)).

Photoaffinity Labeling of *eIF-4E*. To demonstrate saturation effects, samples containing 4 μ g of *eIF-4E* in photolysis buffer (10 mM Tris•Cl, no salt, no DTT, pH 7.6), were incubated at 4°C in 1.5 ml Eppendorf tubes with the appropriate concentration of [γ -³²P]8-N₃GTP for 60 seconds. The reaction mixture was then irradiated for 90 s from a distance of 1 cm with a hand held 254 nm UVB UV lamp (intensity, 1400 μ W/cm²). The total reaction volume was 50 μ l. The reaction was quenched by the addition of 250 μ l of cold acetone. The sample was kept at 4°C for 3 hours and then centrifuged in a Savant HSC10K high speed centrifuge at 10,000 rpm for 20 minutes. The supernatant was drawn off and the pellet was resuspended in a protein solubilizing mixture (PSM) consisting of 10% SDS, 3.6 M urea, 2.5% (w/v) DTT, 2% (w/v) pyronin Y (tracking dye), and 20 mM Tris•Cl (pH 8.0). For studies of protection against photoinsertion, 4 μ g of *eIF-4E* were incubated in photolysis buffer for 60 s at 4°C with the required competitor. At 60 s 40 μ M [γ -³²P]8-N₃GTP was added and at 60 s the samples were irradiated, precipitated, and solubilized as described above.

SDS-PAGE, Scintillation, and Radioisotopic Imaging and Quantitation. Following solubilization, photolabeled samples were analyzed by SDS-PAGE according to Laemmli (Laemmli, U.K. *Nature* 227:680-685 (1970)). The gels were fixed in 25% isopropanol 10% acetic acid for one hour with frequent changes of fixing solution and dried at 80°C on a Bio-Rad Model 583 slab gel dryer. ³²P incorporation into *eIF-4E* was determined with an Ambis 4000 radioisotopic imaging and quantitation system (Scanalytics/CSPI, Billerica, MA).

Photoaffinity Labeling and Enzymatic Digestion of *eIF-4E* for Binding Domain Peptide Isolation. *eIF-4E* (5 mg) was photolabeled with 50 μ M [γ -³²P]8-N₃GTP in a 4 ml total volume of photolysis buffer. The reaction mixture was incubated at 4°C for 5 minutes in a plastic weighing boat followed by irradiation for 2 minutes. This was followed by a second incubation and irradiation with non-radioactive 8-N₃GTP under identical conditions. The reaction was quenched and the protein precipitated by the addition of 4 ml of cold acetone followed by incubation at 4°C for 3 hours. The sample was collected by centrifugation for 20 minutes at 10,000 rpm. The pellet was resuspended in 2 ml of 2 M urea in 75 mM NH₄HCO₃ and the pH adjusted to 8.5-9.0 by the addition of concentrated NH₄OH. *eIF-4E* was proteolyzed by the addition of modified trypsin (10% w/v), with shaking overnight at 20°C.

Immobilized Aluminum(III)-Chelate Chromatography. Iminodiacetic acid-epoxy-activated Sepharose (2 ml) was placed into a 15 ml centrifuge tube and washed as follows: 3 x 10 ml distilled water, 4 x 10 ml 50 mM AlCl_3 , 3 x 10 ml distilled water, 3 x 10 ml buffer A (50 mM NH_4OAC , pH 6.0), 3 x 10 ml buffer B (100 mM NH_4OAC and 0.5 M NaCl , pH 7.0), 3 x 10 ml buffer A. The resin was then transferred to a 10 ml disposable column (Bio-Rad). The labeled eIF-4E digestion mixture was brought to a total volume of 3 ml with buffer A, the pH was adjusted to 6.0 with concentrated acetic acid, and the sample was loaded onto the column. To increase the recovery of the labeled peptide the sample tube was washed with 2 ml buffer A. The column was successively eluted with 2 x 10 ml buffer A, 1 x 10 ml buffer B, 1 x 10 ml buffer A, and the radiolabeled peptide was eluted with 6 ml buffer C (10 mM K_2HPO_4 , pH 8.0). Chromatography was conducted at room temperature with a flow rate of 0.5 ml/min. All fractions were assayed for radioactivity on a Wallac Model 1409 Liquid Scintillation Counter (99% counting efficiency for ^{32}P).

Reverse Phase High Performance Liquid Chromatography and Amino Acid Sequencing of Photolabeled Peptide. The radioactive fractions resulting from buffer C elution from the aluminum(III)-chelate chromatography were pooled and concentrated in a Savant SVC100H speed vac concentrator until a volume of 1.5 ml was reached. 250 μl samples were then analyzed by reversed-phase HPLC using a microbore C_8 column (Brownlee Lab) and an Applied Biosystems 130A separation system. The mobile system consisted of a 25 μM AlCl_3 /0.1% TFA solution (X) and a 0.1% TFA/70% acetonitrile (Y) solvent system. The gradient was as follows; 100% X at a flow rate of 50 $\mu\text{l}/\text{min}$ for 5 minutes, 0-100% Y at a flow rate of 125 $\mu\text{l}/\text{min}$ for 45 minutes, and 100% Y at a flow rate of 125 $\mu\text{l}/\text{min}$ for 5 minutes. Photolabeled peptides were detected by exact correlation of absorbance at 220 nm and ^{32}P cpm. Fractions containing photolabeled peptides were concentrated to 20 μl and subjected to amino acid sequence analysis at the Hunter College Sequencing and Synthesis Facility using an Applied Biosystems 477A protein sequencer and an Applied Biosystems 120A analyzer with on-line PTH derivative identification.

Results

Photoaffinity Labeling of eIF-4E. The specificity of $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ was demonstrated by saturation effects with the probe and by protection studies with m^7GTP , unmethylated nucleotides, and a capped mRNA. Photoinsertion was saturated at $70\text{ }\mu\text{M}$ with an apparent K_d of about $19\text{ }\mu\text{M}$ (Fig. 1A). Increasing concentrations of m^7GTP resulted in decreased photoinsertion by $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ with more than 75% maximal protection observed at $15\text{ }\mu\text{M}$ m^7GTP with an apparent K_d of $7\text{ }\mu\text{M}$ (Fig. 1B). The percent maximal protection of 12% is determined by dividing the $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ concentration ($40\text{ }\mu\text{M}$) by the total maximum nucleotide plus probe concentration ($340\text{ }\mu\text{M}$). This result provides evidence that photoinsertion of $[\gamma\text{-}^{32}\text{P}]\text{8-N}_3\text{GTP}$ was indeed occurring at the m^7G cap binding site of eIF-4E.

To further study the specificity of photoinsertion of the probe into the m^7G mRNA cap binding site the effect of other nucleotides and a capped mRNA was studied (Table 1).

-21--

Table 1. Percent of photolabeling of ,eIF-4E by 40 μ M [γ - 32 P]8N₃GTP in the presence of various nucleotides and capped mRNA

	Nucleotide	nucleotide concentration	% of control
5	none	0 μ M	100
	GTP	300 μ M	21
	GDP	300 μ M	26
	GMP	300 μ M	73
	CTP	300 μ M	43
10	ATP	300 μ M	35
	ADP	300 μ M	51
	AMP	300 μ M	89
	UTP	300 μ M	58
	UDP	300 μ M	93
15	UMP	300 μ M	93
	m ⁷ GTP	300 μ M	16
	capped mRNA ^a	300 μ M	5

^acapped mRNA, 34 base long oligoribonucleotide

Among the unmethylated nucleotides, GTP (21% of control), GDP (26% of control), and ATP (36% of control) proved to be the best inhibitors to photoinsertion. Although other nucleotides were also able to act as inhibitors they were much less effective. Capped mRNA was able to dramatically reduce photoinsertion (5% of control) as compared to equimolar concentrations of m⁷GTP (18% of control) and unmethylated nucleotides (maximum of 21% of control).

Isolation and Identification of [γ - 32 P]8-N₃GTP Photolabeled Peptides. ,eIF-4E was photolabeled as described above to determine the amino acid sequence that was covalently modified by [γ - 32 P]8-N₃GTP within the m⁷G cap binding site. The second round of photolysis was performed with nonradioactive 8-N₃GTP to increase the amount of binding domain peptide while reducing nonspecific radiolabeling. TCA was initially used to precipitate photolabeled ,eIF-4E, however this approach resulted in low yields of peptide for sequencing probably due to the

disruption of the acid sensitive probe-peptide bond. However, acetone precipitation of photolabeled eIF-4E resulted in higher yields of labeled peptide.

Photolabeled eIF-4E was digested overnight with modified trypsin. The labeled peptide was isolated by affinity chromatography and further purified by reversed-phase HPLC, using the methods described above. To locate the photolabeled peptide in the various buffer solutions the flow through fractions (fractions 1-5), washes (fractions 6-25), and phosphate elutions (fractions 26-31) (Fig. 2) were separately pooled and subjected to reversed-phase HPLC. The flow through and wash fractions contained no radioactive peaks as determined by liquid scintillation counting thus no labeled peptides. However, these fractions did contain many peptide fragments, demonstrating that most of the peptides are unmodified and not retained on the Al³⁺ resin. The HPLC absorbance profile at 220 nm and the corresponding ³²P cpm profile for the phosphate elution (fractions 16-31) is shown in figure 3. Two major radioactive peaks were observed which also gave a UV absorbance. The first peak at 3 minutes represents the flow through and injection disturbance and contained no peptides as revealed by sequencing data. The radioactivity detected with this peak was free photolyzed [γ -³²P]8-N₃GTP, or probe hydrolyzed during HPLC due to the lability of the N-glycosyl bond to acidic conditions (King, S., Kim, H., & Haley, B., *Methods Enzymol.* 196:449-466 (1991)). The second major radioactive peak at 23 minutes was concentrated for amino acid sequence analysis. Photolabeling, peptide isolation, and sequencing were repeated in three separate experiments to ensure that only one binding domain was covalently modified by [γ -³²P]8-N₃GTP. In all experiments the radioactivity coeluted with fraction 23. The sequenced peptide consisted of residues 113-122 of human eIF-4E which includes tryptophan 6 (see Table 2).

Table 2. Sequence analysis of [γ - ^{32}P]8N₃GTP the photolabeled tryptic peptide from the radioactive peak at 23 minutes in the HPLC chromatogram

	cycle number	identified residue	picomoles observed
5	1	W	5.6
	2	L	17.5
	3	I	14.1
	4	T	10.6
	5	L	11.7
10	6	N	7.0
	7	K	NF ^a
	8	Q	6.4
	9	Q	4.9
	10	R	2.1
15	^a NF, not found		

In addition the lysine residue at position 119 remained uncleaved by trypsin. This information combined with the lack of identification of lysine 119 by sequence analysis indicates that the residue is the likely site of photoinsertion. No minor peptides were detected in the sequencing data.

20 Example 2: Binding of Capped mRNA to Synthetic Peptides

To more closely examine the eIF-4E cap-binding domain, synthetic peptides were produced and the binding of m⁷GTP to these peptides were measured. With one peptide (WKRGGRWLITLNKQQR; SEQ ID NO:7), direct fluorescence titration experiments were used to measure the equilibrium binding constant for the formation of the 4E/m⁷GTP complex.

25 Measurements were made in 20 mM HEPES.KOH/1mM DTT, pH 7.6, at 25°C on a SPEX fluorolog-T2 spectrofluorometer equipped with a high-intensity (450 W) xenon arc lamp.

Changes in fluorescence (ΔF) over time were plotted on an Eadie-Hofstee plot. In these experiments, an equilibrium binding constant of 8.5×10^7 was calculated from the plotted results.

Analogous experiments were conducted with a second peptide (WEDEKNKRGGRWLITLNKQKK; SEQ ID NO:8), using capped mRNA at a ratio of 1:20 (RNA:peptide). Binding of capped mRNA to the peptide was determined in gel shift assays under various ionic and buffer conditions. In a binding solution of 20 mM HEPES, pH 7.6, 1 mM DTT, 50mM NaCl and 5% glycerol, there was a shift of the peptide/mRNA complex towards the negative pole of the gel. When binding was conducted in 10 mM potassium phosphate, pH 8.0, 0.25 M KCl, 2 mM EDTA, 5% glycerol, however, an electrophoretic shift was observed towards the positive pole. The net charge of the peptide in the presence and absence of phosphate in these experiments was 4^- and 3^+ , respectively.

Discussion

The present experiments demonstrate the ability of $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ to serve as a substitute for the $m^7(5')\text{Gppp}(5')\text{N}$ cap structure of mRNA in binding to eIF-4E. Selectivity of the probe for the active site was determined in saturation studies with $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ and in protection studies utilizing $m^7\text{GTP}$, various nucleotides, and capped mRNA as inhibitors to photoinsertion. Photoinsertion into eIF-4E was saturated with $70\text{ }\mu\text{M}$ $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ and the apparent K_d was $19\text{ }\mu\text{M}$. The observation that photoinsertion of $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ was inhibited by 90% in the presence of $90\text{ }\mu\text{M}$ $m^7\text{GTP}$ demonstrates the specificity of this binding site. The ability of the purine nucleotides GTP, GDP, and ATP to reduce photoinsertion is not surprising because the cap binding site is a dinucleotide site and the probe used in these studies was a mononucleotide. Additional evidence for the site specific insertion of $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ into the cap binding domain is that of capped mRNA ($m^7\text{G}(5')\text{ppp}(5')\text{G}$, 34 nucleotides) as the inhibitor which drastically reduced photoinsertion of the probe to 5% of the control (no inhibitor).

The peptide resulting from photoinsertion of $[\gamma\text{-}^{32}\text{P}]8\text{-N}_3\text{GTP}$ and subsequent isolation has been identified as the tryptic peptide consisting of residues 113-122 and thus contains tryptophan 6. Detection of the actual site of nitrene insertion is often difficult due to the lability

of the N-glycosyl bond to the acidic conditions of HPLC (King, S., Kim, H., & Haley, B., *Meth. Enzymol.* 196:449-466 (1991)). However these data suggest that lysine 119 is the modified residue for two reasons: the photolabeled binding site peptide is the product of tryptic cleavage at Arg-112 and Arg-122 yet Lys-119 remained uncleaved; and amino acid sequence analysis was
5 unable to identify Lys-119 despite the continuation of sequencing for three residues beyond this position. Such a gap in sequencing usually indicates reaction of the probe with the missing amino acid (Shoemaker, M.T. & Haley, B.E., *Biochemistry* 32:1883-1890 (1993)). The amino acid is "missing" because it does not elute in the expected position when the probe has been covalently attached to it.

10 Fluorescence and circular dichroism studies have been performed with Trp→Lys mutants of tryptophan 6 of yeast eIF-4E which indicate that this residue may be responsible for imparting specificity to cap recognition (McCubbin, W.D., *et al.*, *J. Biol. Chem.* 263:17663-17671 (1988)). It was found that 7-methylated analogs of GDP and G(5')ppp(5')G bind strongly to the mutant, but unmethylated analogs also have binding affinity. The presence of the phosphate groups results
15 in a stacking interaction with the indole ring regardless of the state of methylation. This has been attributed to electrostatic or hydrogen-bonding between the tryptophan side chain and the highly charged phosphates (Kamiichi, K., *et al.*, *J. Chem. Soc. Perkin Trans. II*, 1739-1745 (1987)). The substitution of glycine 111 by aspartic acid in yeast eIF-4E reduced cap binding activity as measured by binding of eIF-4E to m⁷GDP agarose affinity columns (Altmann, M. & Trachsel, H.,
20 *Nucleic Acids Res.* 9:1643-1656 (1989)). This glycine is two residues to the N-terminal side of tryptophan 6.

The identification of residues 113-122 as playing a role in m⁷GpppG cap recognition by eIF-4E is consistent with previous studies. The eight tryptophans in yeast eIF-4E have been classified into three separate groups on the basis of the mRNA cross-linking activity of Trp→Phe
25 mutants: 1) tryptophans that are strongly required for eIF-4E cross-linking to the mRNA cap which include tryptophans 1, 2, 5 and 8; 2) tryptophans that reduce the cross-linking ability of eIF-4E which include 3, 6 and 7; 3) tryptophans that are not required for cap recognition which includes 4. Furthermore, the mutation of tryptophans 1 and 8 completely obliterated cap binding activity (Altmann, M., *et al.*, *J. Biol. Chem.* 263:17229-17232 (1988)). The large number of
30 residues affecting binding may indicate that the overall folding of eIF-4E plays a crucial role in

cap recognition. Identification of amino acid residues that participate in the m⁷G(5')ppp(5)N binding site of eIF-4E will aid in further mutational studies designed to understand the molecular details of eIF-4E•cap recognition. In addition, further knowledge of the structural motif required for eIF-4E•mRNA cap interactions will aid our understanding of precisely how translational repressor 4E-binding proteins function (Lin, T.A., *et al. Science* 266, 653-656 (1994); Mader, S., *et al., Mol. Cell Biol.* 15:4990-4997 (1995)).

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

-27- .

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- 5 (A) NAME: Life Technologies, Inc.
(B) STREET: 9800 Medical Center Drive
(C) CITY: Rockville
(D) STATE: Maryland
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 20850

- 10 (ii) TITLE OF INVENTION: Cap-Binding Domain of Human Eukaryotic Protein
Synthesis Initiation Factor eIF-4E and the Use Thereof

(iii) NUMBER OF SEQUENCES: 27

(iv) COMPUTER READABLE FORM:

- 15 (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) PRIOR APPLICATION DATA:

- 20 (A) APPLICATION NUMBER: US (to be assigned)
(B) FILING DATE: 29-AUG-1997

(vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/024,881
(B) FILING DATE: 30-AUG-1996

25 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg
1 5 10

35 (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids

-28-

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 11..13

10 (D) OTHER INFORMATION: /note= "residues at positions 11 and 13 may be charged, aromatic or beta-turn-inducing amino acids (such as D, E, Y, T, G or P); residue at position 12 may be a hydrophobic amino acid (such as L, I or V)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Xaa Xaa Xaa
1 5 10

15 (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

20 (C) STRANDEDNESS: single

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 1

25 (D) OTHER INFORMATION: /note= "residue at position 1 may be any amino acid"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 12

30 (D) OTHER INFORMATION: /note= "residue at position 12 may be any amino acid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Xaa Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Xaa
1 5 10

35 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

-29-

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /note= "residue at position 1 may be any amino acid"

10

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 18
- (D) OTHER INFORMATION: /note= "residue at position 18 may be any amino acid"

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Xaa	Trp	Lys	Arg	Gly	Gly	Arg	Trp	Leu	Ile	Thr	Leu	Asn	Lys	Gln	Gln
1				5				10						15	

Arg Xaa

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

25

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /note= "residue at position 1 may be any amino acid"

30

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 23
- (D) OTHER INFORMATION: /note= "residue at position 23 may be any amino acid"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

-30-

Xaa Trp Glu Asp Glu Lys Asn Lys Arg Gly Gly Arg Trp Leu Ile Thr
 1 5 10 15

Leu Asn Lys Gln Gln Arg Xaa
 20

5 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
 (B) LOCATION: 1..3
 (D) OTHER INFORMATION: /note= "residues at positions 1 and 3 may be charged, aromatic or beta-turn-inducing residues (such as D, E, Y, T, G or P); residue at position 2 may be a hydrophobic amino acid (such as L, I, V or W)"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
 (B) LOCATION: 14..16
 (D) OTHER INFORMATION: /note= "residues at positions 14 and 16 may be charged, aromatic or beta-turn-inducing amino acids (such as D, E, Y, T, G or P); residue at position 15 may be a hydrophobic amino acid (such as L, I, V or W)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Xaa Xaa Xaa Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Xaa Xaa Xaa
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

-31-

Trp Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

- 10 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Trp Glu Asp Glu Lys Asn Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu Asn Lys
 1 5 10 15
 Gln Gln Lys Lys
 20

(2) INFORMATION FOR SEQ ID NO:9:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

25 Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Gly Asp Ile Val Val Leu
 1 5 10 15
 Val Pro

(2) INFORMATION FOR SEQ ID NO:10:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

- (ii) MOLECULE TYPE: peptide

-32- .

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Asp Gly Val Val Ile Ile
 1 5 10 15

Val Leu Leu Asp Gly
 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Asp Leu Leu Val Ile Val
 1 5 10 15

Glu Gly

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Glu Asp Val Ile Val Leu
 1 5 10 15

Leu Ile Tyr

(2) INFORMATION FOR SEQ ID NO:13:

-33-

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Pro Ile Val Val Ile Val
1 5 10 15

10 Leu Val Ile Ile Asp Gly
20

(2) INFORMATION FOR SEQ ID NO:14:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

20 Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Asp Thr Val Leu Val Val
1 5 10 15

Ile Gly Pro

(2) INFORMATION FOR SEQ ID NO:15:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

30 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

-34-

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Gly Asp Ile Ile Val Val
 1 5 10 15

Val Ile Val Asp Pro
 20

5 (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Gly Val Val Val Ile Ile
 1 5 10 15

Val Leu Leu Asp
 20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Pro Thr Val Ile Val Val
 1 5 10 15

Ile Asp

30 (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single

-35-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

5 Trp Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

15 Trp Glu Asp Glu Lys Asn Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu
 1 5 10 15

Asn Lys Gln Gln Lys Lys
 20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 15 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

25 Lys Leu Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Val Leu Gln
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

-36-

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ile Gln Leu Val Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Leu Lys
1 5 10 15

Ala Gly

10

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

15

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Val Ala Trp Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu Ile Thr Leu
1 5 10 15

20

Asn Lys Gln Gln Arg Gln Ala Lys Thr
20 25

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

25

(ii) MOLECULE TYPE: peptide

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

-37-

Ile Thr Val Gln Ala Trp Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu
 1 5 10 15

Asn Lys Gln Gln Arg Lys Thr Leu Phe
 20 25

5 (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Leu Arg Val Trp Glu Asp Glu Lys Asn Lys Arg Gly Gly Arg Trp Leu
 1 5 10 15

15

Ile Thr Leu Asn Lys Gln Gln Arg Lys Tyr
 20 25

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

20

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

25

Asn Ile Ser Trp Lys Arg Gly Gly Arg Trp Leu Ile Thr Leu Asn Lys
 1 5 10 15

Gln Gln Arg Phe Lys Ala Val Gln Gln
 20 25

(2) INFORMATION FOR SEQ ID NO:26:

30

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

-38-

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Glu Tyr Glu Gly Val Val Pro Trp Leu Ile Thr Leu Asn Lys Gly Glu
1 5 10 15

5 Arg Asp Leu Ile Leu Val Ile Gly
20

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

15 Asp Gly Pro Gly Leu Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Tyr
1 5 10 15

Thr Trp Val Val Ile Leu Val Thr Gly
20 25

WHAT IS CLAIMED IS:

1. A truncated peptide of the cap-binding protein consisting essentially of the mRNA cap-binding domain.
2. The peptide of claim 1, wherein said cap-binding protein is eIF-4E.
- 5 3. The peptide of claim 1, wherein said peptide has an amino acid sequence corresponding to SEQ ID NO:1.
4. The peptide of claim 3, wherein said peptide contains one or more substitutions, deletions and/or deletions and wherein said peptide maintains mRNA cap binding function.
5. The peptide of claim 1, wherein said peptide has an amino acid sequence
 10 corresponding to a general pattern selected from the group consisting of Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X₁₍₁₋₂₎-X₂₍₄₋₁₀₎-X₁₍₁₋₂₎ (SEQ ID NO:2), and X₁₍₀₋₁₀₎-X₂₍₀₋₂₎-X₁₍₀₋₁₀₎-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X₁₍₁₋₂₎-X₂₍₄₋₁₀₎-X₁₍₁₋₁₀₎ (SEQ ID NO:6),
 wherein X₁ is a charged, aromatic or β -turn-inducing amino acid; X₂ is a hydrophobic amino acid; and wherein the subscripts for X₁ and X₂ indicate the number of such
 15 residues in consecutive sequence at the indicated position.
6. The peptide of claim 1, wherein said peptide has an amino acid sequence corresponding to the group consisting of:
 - (a) X_n-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X_n (SEQ ID NO:3);
 - (b) X_n-Trp-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-
 20 X_n (SEQ ID NO:4); and
 - (c) X_n-Trp-Glu-Asp-Glu-Lys-Asn-Lys-Arg-Gly-Gly-Arg-Trp-Leu-Ile-Thr-Leu-Asn-Lys-Gln-Gln-Arg-X_n (SEQ ID NO:5),
 wherein X is any amino acid, and n is a whole number ranging in value from 0 to 70.

7. The peptide of claim 5, wherein X1 is Asp, Glu, Tyr, Thr, Gly or Pro.
8. The peptide of claim 5, wherein X2 is Leu, Ile, Val or Trp.
9. The peptide of claim 6, wherein n is a whole number ranging in value from 0 to 25.
- 5 10. The peptide of claim 6, wherein n is a whole number ranging in value from 0 to 10.
11. The peptide of claim 1, wherein said peptide has an amino acid sequence corresponding to SEQ ID NO:7.
12. The peptide of claim 1, wherein said peptide has an amino acid sequence
10 corresponding to SEQ ID NO:8.
- 13 The peptide of any one of claims 1-6, 11 or 12, wherein said peptide is immobilized on a solid support.
14. The peptide of claim 13, wherein said peptide is immobilized by an amino acid linkage group or a chemical linkage.
- 15 15. The peptide of claim 14, wherein one end of the peptide is immobilized to the solid support.
16. The peptide of claim 13, wherein said solid support is selected from the group consisting of nitrocellulose, diazocellulose, glass, latex, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, silica gel, a bead and a
20 microtitre plate.

17. The peptide of claim 13 wherein said solid support is a glass bead, a latex bead or a magnetic bead.

18. A resin comprising the peptide of claim 1 covalently or non-covalently attached to a solid support matrix.

5 19. The resin according to claim 18, wherein said attachment is accomplished by an amino acid linking group or a chemical linking group.

20. The resin of according to claim 18, wherein one end of the peptide is covalently or non-covalently attached to the resin.

10 21. The resin of claim 18, which is selected from the group consisting of nitrocellulose, diazocellulose, glass, latex, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, silica gel and a bead.

22. The resin of claim 18, which is a glass bead, a latex bead or a magnetic bead.

23. A method for the isolation of a capped mRNA molecule, said method comprising the steps of

15 (a) contacting a solution containing said capped mRNA molecule with the peptide of any one of claims 1-6, 11 or 12 under conditions favoring the noncovalent binding of said capped mRNA molecule to said peptide; and

(b) removing materials not bound to said peptide.

20 24. The method of claim 23, further comprising an elution step (c), wherein said capped mRNA molecule is eluted from said peptide using an elution buffer comprising a competitive analog of the cap region of a mRNA molecule, a changed the pH, or changed ionic strength.

25. The method of claim 23, wherein said peptide is immobilized on a solid support.

26. The method of claim 25, wherein said solid support is selected from the group consisting of nitrocellulose, diazocellulose, glass, latex, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, silica gel, a bead and a
5 microtitre plate.

27. The method of claim 26, wherein said bead is a glass bead, a latex bead or a magnetic bead.

28. A method for the production of a full-length cDNA molecule, said method comprising the steps of

- 10 (a) contacting a solution containing a capped mRNA molecule with a protein having reverse transcriptase activity under conditions favoring the production of a plurality of cDNA molecules from said mRNA and the formation of a mixture of mRNA:cDNA hybrids;
- (b) contacting said mixture of mRNA:cDNA hybrids with a single-stranded RNA-specific nuclease under conditions favoring the digestion of RNA not hybridized to cDNA;
- 15 (c) contacting said mixture of mRNA:cDNA hybrids with an endonuclease that specifically attacks the 3' adjacent phosphodiester-bound GpN;
- (d) contacting said mixture of mRNA:cDNA hybrids with the peptide of any one of claims 1-6, 11 or 12, under conditions favoring the noncovalent binding to said peptide of said mRNA:cDNA hybrids;
- 20 (e) removing materials not bound to said peptide; and
- (f) eluting said bound mRNA:cDNA hybrids from said peptide.

29. The method of claim 28, wherein said protein having reverse transcriptase activity is a truncated or mutated reverse transcriptase enzyme which lacks RNase H activity.

30. The method of claim 28, wherein said protein having reverse transcriptase activity
25 is a DNA polymerase having reverse transcriptase activity.

31. The method of claim 30, wherein said DNA polymerase is a thermostable DNA polymerase.

32. The method of claim 31, wherein said DNA polymerase is selected from the group consisting of *Tne*, *Tth*, and *Taq* or a mutant thereof.

5 33. The method of claim 28, wherein said single-stranded RNA-specific nuclease is T_1 nuclease.

34. The method of claim 28, wherein said elution step (f) is accomplished using a competitive analog of the cap region of a mRNA molecule, changing the pH, or the ionic strength of the solution.

10 35. The method of claim 28, wherein said peptide is immobilized on a solid support.

36. A kit for the isolation of a capped mRNA molecule, said kit comprising one or more containers, wherein a first container contains the peptide of any one of claims 1-6, 11 or 12, and wherein a second container contains an elution buffer for recovering the capped mRNA from said peptide.

15 37. A kit for the production of a full-length cDNA molecule, said kit comprising one or more containers, wherein a first container contains the peptide of any one of claims 1-6, 11 or 12, a second container contains a protein having reverse transcriptase activity, and a third container contains a nuclease enzyme.

38. The kit of claim 36, wherein said peptide is immobilized on a solid support.

20 39. The kit of claim 37, wherein said peptide is immobilized on a solid support.

40 The kit of claim 37, wherein said protein having reverse transcriptase activity is a truncated or mutated reverse transcriptase enzyme lacking RNase H activity.

41. The kit of claim 37, wherein said protein having reverse transcriptase activity is a DNA polymerase having reverse transcriptase activity.

5 42. The kit of claim 37, wherein said nuclease enzyme is T_1 nuclease.

1/3

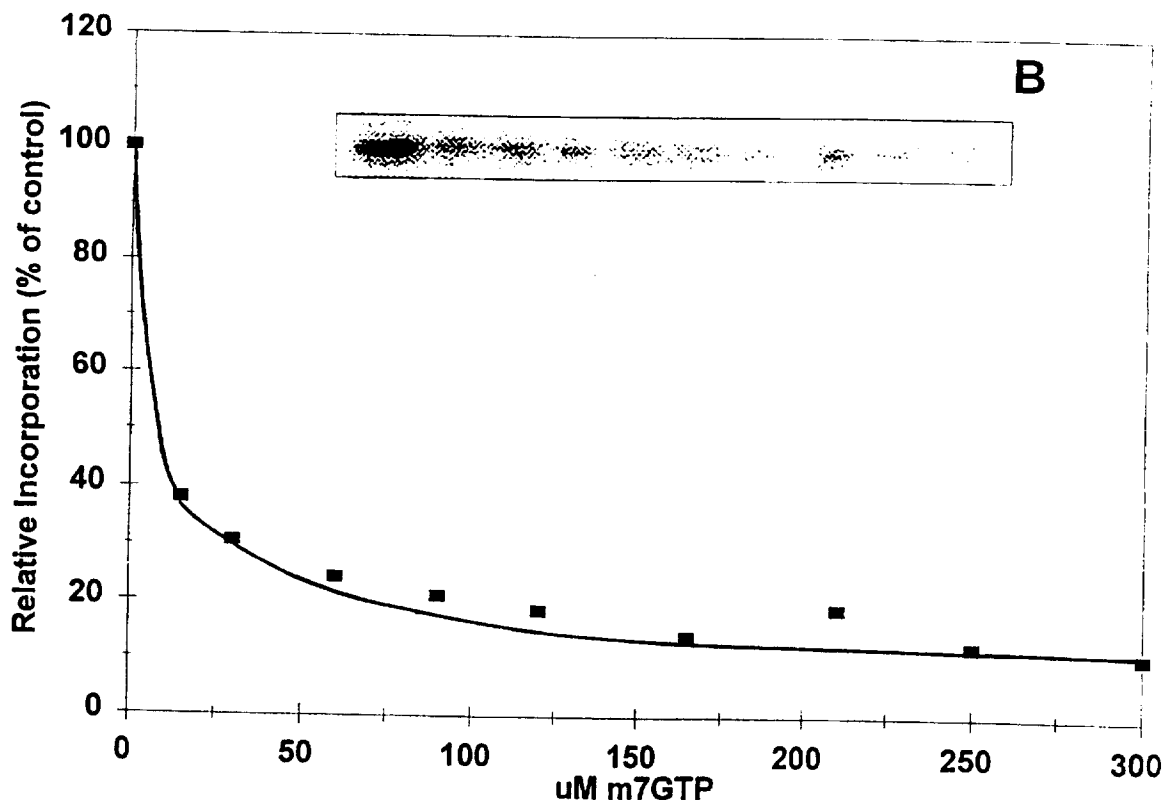
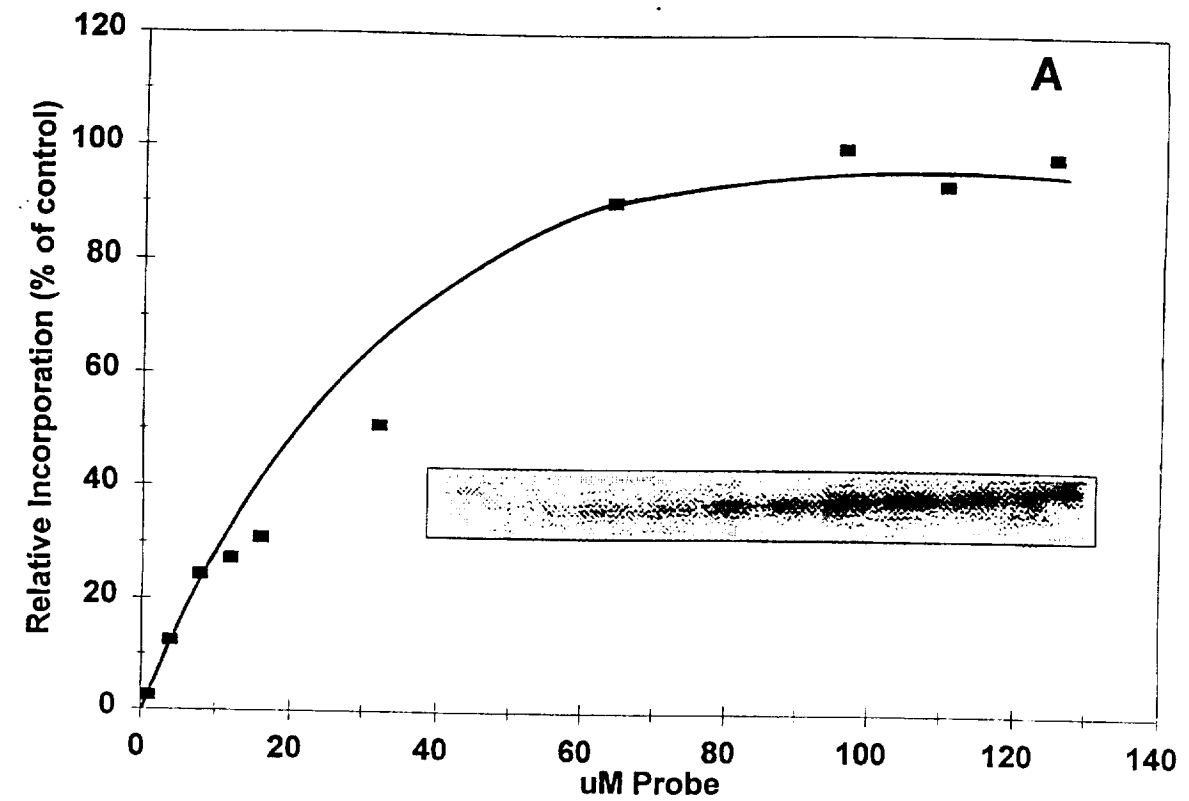


figure 1

2/3

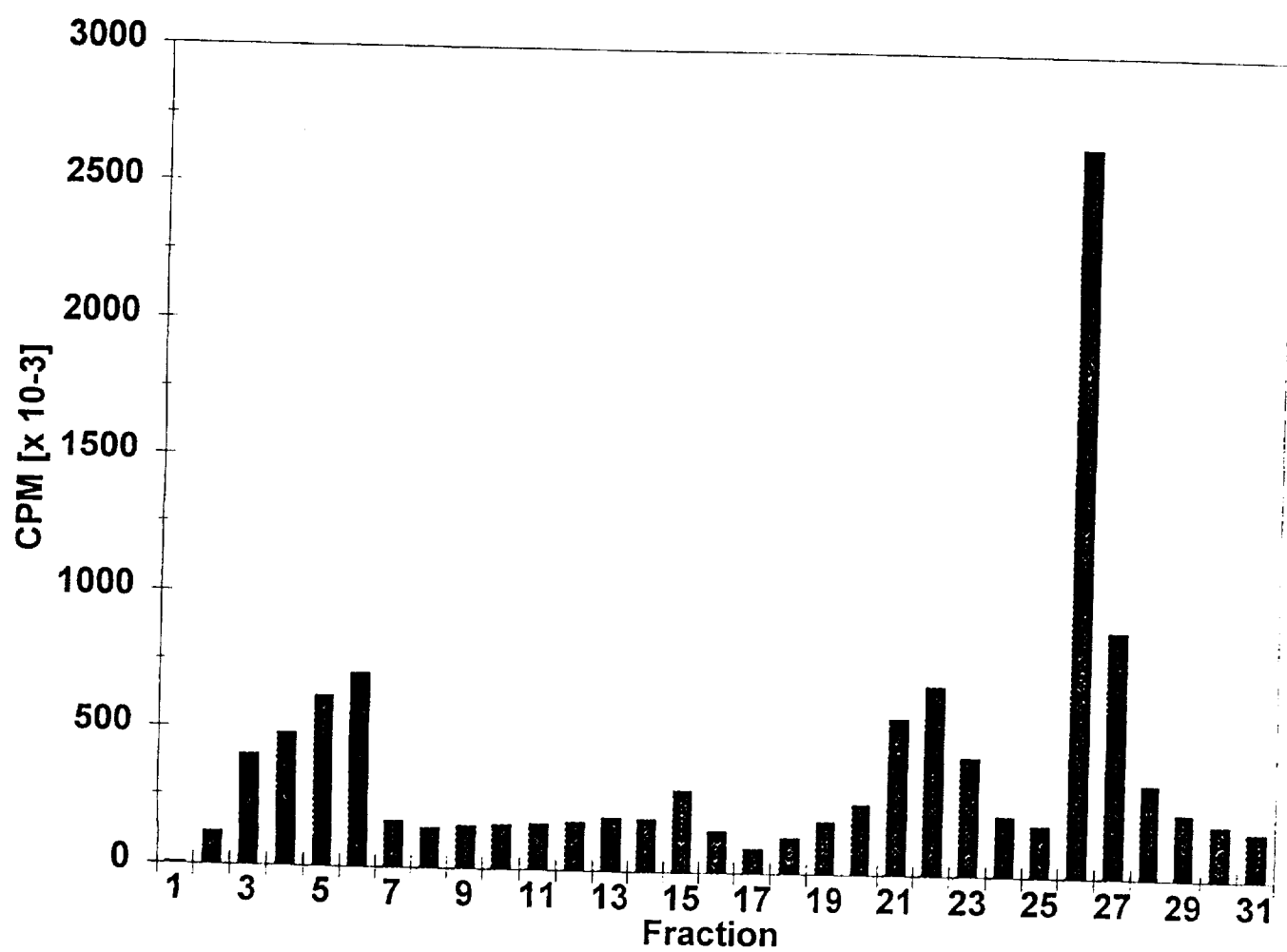


figure 2

3/3.

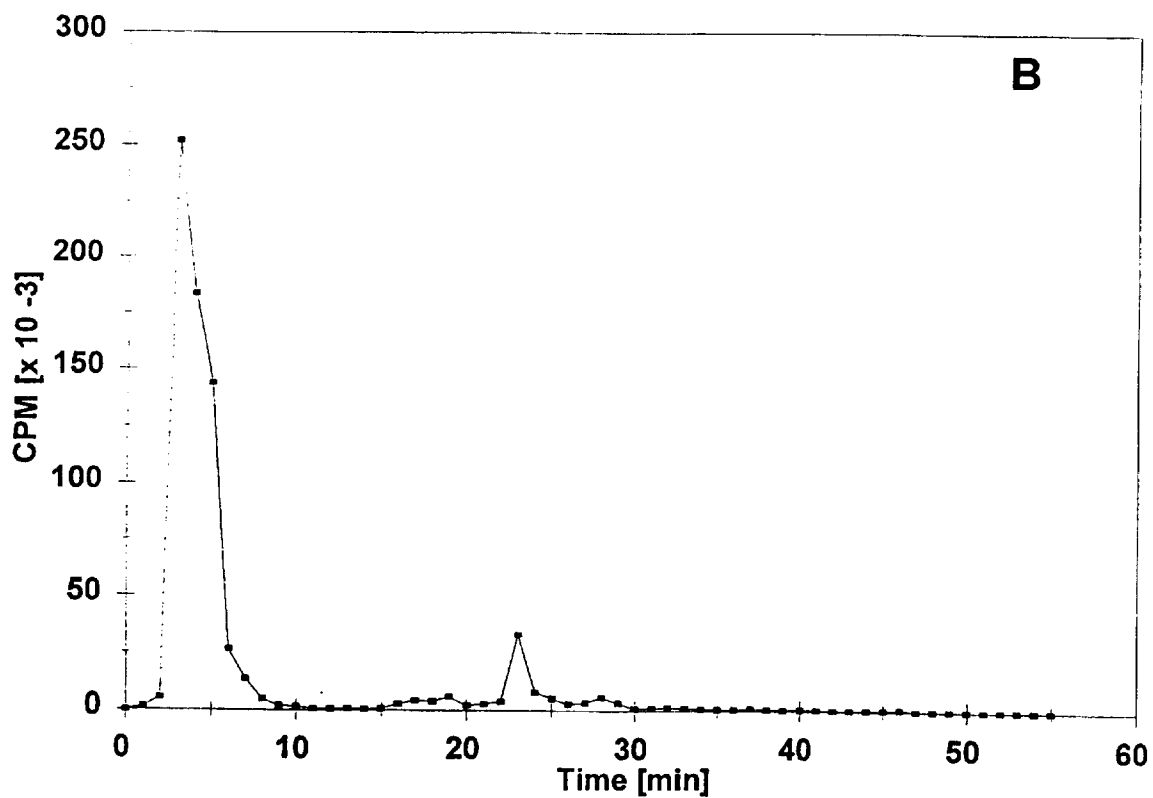
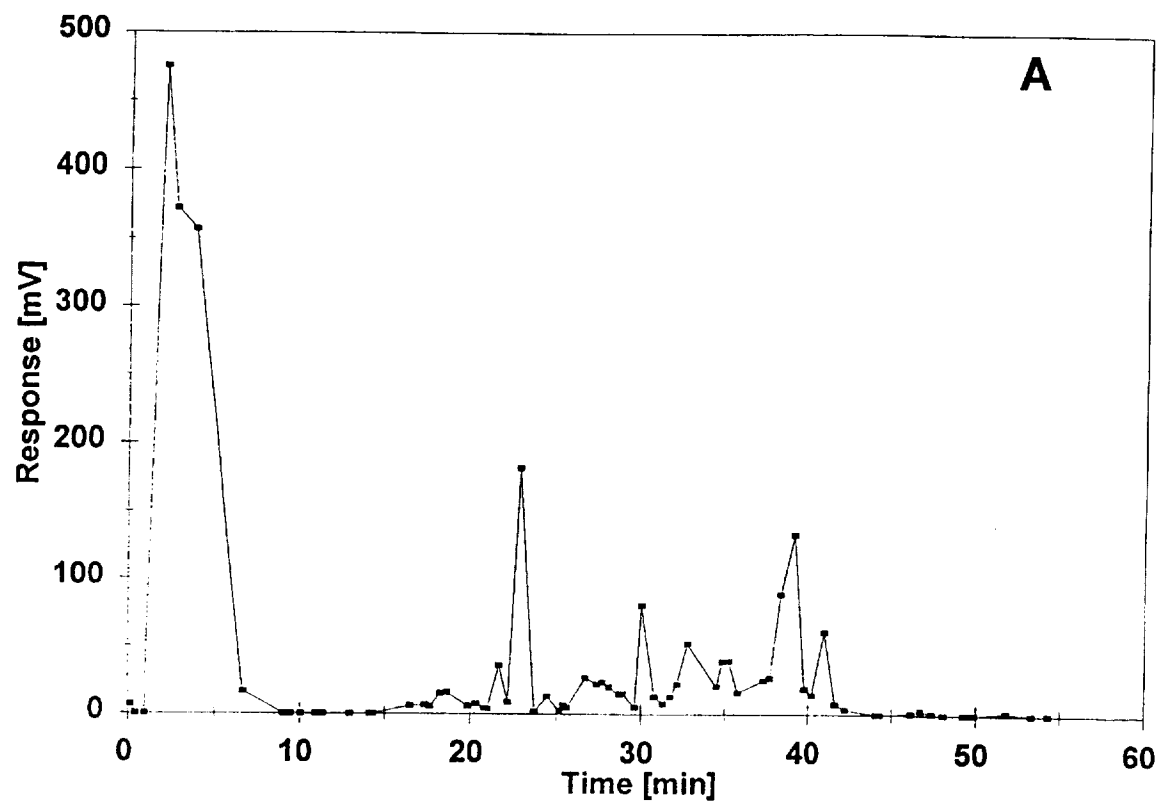


figure 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/15295

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 21/04; C07K 1/04, 1/14, 1/36, 2/00, 4/00, 5/02, 7/02

US CL :530/300, 324, 325, 326, 327, 328, 329, 330, 333, 334; 536/25.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/300, 324, 325, 326, 327, 328, 329, 330, 333, 334; 536/25.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS, BIOSIS, DERWENT

search terms: eif-4e, cap-bind? protein?, mRNA cap-bind?, immobil?, attach? bound?, solid? support?, bead?, resin?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MORINO et al. Direct Expression of a Synthetic Gene in Escherichia coli: Purification and Physicochemical Properties of Human Initiation Factor 4E. J. Biochem. 1994, Vol. 116, No. 3, pages 687-693, especially the abstract, Figure 1 and Figure 2.	1-42
Y	RYCHLIK et al. Amino Acid Sequence of the mRNA Cap-Binding Protein from Human Tissues. Proc. Natl. Acad. Sci. February 1987, Vol. 84, No. 4, pages 945-949, especially the abstract, Table 1 and Figure 2.	1-42

☒ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 03 DECEMBER 1997	Date of mailing of the international search report 22 DEC 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer NANCY J. DEGEN Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/15295

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	FRIEDLAND et al. Identification of the Cap Binding Domain of Human Recombinant Eukaryotic Protein Synthesis Initiation Factor 4E Using a Photoaffinity Analogue. Protein Sci. January 1997, Vol. 6, No. 1, pages 125-131, especially the abstract and Table 1.	1-42